

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Construction of Forward-Looking Distributions Using Limited Historical Data and Scenario Assessments

*Riaan de Jongh, Helgard Raubenheimer
and Mentje Gericke*

Abstract

Financial institutions are concerned about various forms of risk that might impact them. The management of these institutions has to demonstrate to shareholders and regulators that they manage these risks in a pro-active way. Often the main risks are caused by excessive claims on insurance policies or losses that occur due to defaults on loan payments or by operations failing. In an attempt to quantify these risks, the estimation of extreme quantiles of loss distributions is of interest. Since financial companies have limited historical data available in order to estimate these extreme quantiles, they often use scenario assessments by experts to augment the historical data by providing a forward-looking view. In this chapter, we will provide an exposition of statistical methods that may be used to combine historical data and scenario assessments in order to estimate extreme quantiles. In particular, we will illustrate their use by means of practical examples. This method has been implemented by major international banks and based on what we have learnt in the process, we include some practical suggestions for implementing the recommended method.

Keywords: operational risk, loss distribution approach, aggregate loss distribution, historical data, measures of agreement, scenario assessments

1. Introduction

Financial institutions need to carefully manage financial losses. For example, the claims made against short-term insurance policies need to be analysed in order to enable an insurance company to determine the reserves needed to meet their obligations and to determine the adequacy of their pricing strategies. Similarly, banks are required in terms of regulation to set aside risk capital to absorb unexpected losses that may occur. Of course, financial institutions are more interested in the total amount of claims or the aggregate loss occurring over one year in the future, than the individual claims or losses. For this reason, their focus will be on what may happen in the year ahead rather than what has happened in the past. Popular modelling methods involve the construction of annual aggregate claim or loss distributions using the so-called loss distribution approach (LDA) or random sums

method. Such a distribution is assumed to be an adequate reflection of the past but need to be forward looking in the sense that anticipated future losses are taken into account. The constructed distribution may then be used to answer questions like ‘What aggregate loss level will be exceeded only once in c years?’ or ‘What is the expected annual aggregate loss level?’ or ‘If we want to guard ourselves against a one in a thousand-year aggregate loss, how much capital should we hold next year?’ The aggregate loss distribution and its quantiles will provide answers to these questions and it is therefore paramount that this distribution is modelled and estimated as accurately as possible. Often it is the extreme quantiles of this distribution that is of interest.

Under Basel II’s advanced measurement approach, banks may use their own internal models to calculate their operational risk capital, and the LDA is known to be a popular method for this. A bank must be able to demonstrate that their approach captures potentially severe ‘tail’ events and they must hold capital to protect them against a one-in-a-thousand-year aggregate loss. To determine this capital amount, the 99.9% Value-at-Risk (VaR) of the aggregate distribution is calculated [1]. In order to estimate a one-in-a-thousand-year loss, one would hope that at least a thousand years of historical data is available. However, in reality only between five and ten years of internal data is available and scenario assessments by experts are often used to augment the historical data and to provide a forward-looking view.

The much anticipated implementation of Basel III will require banks to calculate operational risk capital on a new standardised approach, which is simple, risk-sensitive and comparable between different banks [2]. Although the more sophisticated internal models described above will no longer be allowed in determining minimum regulatory capital, these models will remain relevant for the determination of economic capital and decision making within banks and other financial institutions. It is also suggested that LDA models would form an integral part of the supervisory review of a bank’s internal operational risk management process [3]. For this reason, we believe the LDA remains relevant and will continue to be studied and improved on.

In this chapter we provide an exposition of statistical methods that may be used to estimate VaR using historical data in combination with quantile assessments by experts. The proposed approach has been discussed and studied elsewhere (see [4]), but specifically in the context of operational risk and economic capital estimation. In this chapter we concentrate on the estimation of the VaR of the aggregate loss or claims distribution and strive to make the approach more accessible to a wider audience. Also, based on the implementation done for major banks, we include some practical guidelines for the use and implementation of the method in practice. In the next section we discuss two approaches, Monte Carlo and Single Loss Approximation, that may be used for the approximation of VaR assuming known distributions and parameters. Then, in the third section (Historical data and scenario modelling), we will discuss the available sources of data and formulate the scenario approach and how these may be created and assessed by experts. This is followed, in section four (Estimating VaR), by the estimation of VaR using three modelling approaches. In the fifth section (Implementation recommendations) some guidelines on the implementation of the preferred approach are given. Some concluding remarks are made in the last section.

2. Approximating VaR

Let the random variable N denotes the annual number of loss events and that N is distributed according to a Poisson distribution with parameter λ ,

i.e. $N \sim Poi(\lambda)$. Note that one could use other frequency distributions like the negative binomial, but we found that the Poisson is by far the most popular in practice since it fits the data well. Furthermore, assume that the random variables X_1, \dots, X_N denote the loss severities of these loss events and that they are independently and identically distributed according to a severity distribution T , i.e. $X_1, \dots, X_N \sim iid T$. Then the annual aggregate loss is $A = \sum_{n=1}^N X_n$ and the distribution of A is the aggregate loss distribution, which is a compound Poisson distribution that depends on λ and T and is denoted by $CoP(T, \lambda)$. Of course, in practice we do not know T and λ and have to estimate it. First we have to decide on a model for T , which can be a class of distributions $F(x, \theta)$. Then θ and λ have to be estimated using statistical estimates.

The compound Poisson distribution $CoP(T, \lambda)$ and its VaR are difficult to calculate analytically so that in practice Monte Carlo (MC) simulation is often used. This is done by generating N according to the assumed frequency distribution and then by generating X_1, \dots, X_N independent and identically distributed according to the true severity distribution T and calculating $A = \sum_{n=1}^N X_n$. The previous process is repeated I times independently to obtain $A_i, i = 1, 2, \dots, I$ and then the 99.9% VaR is approximated by $A_{([0.999 * I] + 1)}$ where $A_{(i)}$ denotes the i -th order statistic and $[k]$ the largest integer contained in k . Note that three input items are required to perform this, namely the number of repetitions I as well as the frequency and loss severity distributions. The number of repetitions determines the accuracy of the approximation and the larger it is, the higher its accuracy. In order to illustrate the Monte Carlo approximation method, we assume that the Burr is the true underlying severity distribution and we use six parameter sets corresponding to an extreme value index (EVI) of 0.33, 0.83, 1.0, 1.33, 1.85 and 2.35 as indicated in **Table 1** below. See Appendix A for a discussion of the characteristics of this distribution and its properties. We take the number of repetitions as $I = 1\,000\,000$ and repeat the calculation of VaR 1000 times. The 90% band containing the VaR values are shown in **Figure 1** below. Here the lower (upper) bound has been determined as the 5% (95%) percentile of the 1000 VaR values, divided by its median, and by subtracting 1. In mathematical terms the 90% band is defined as $\left[\frac{VaR_{(51)}}{Median(VaR_1, \dots, VaR_{1000})} - 1, \frac{VaR_{(951)}}{Median(VaR_1, \dots, VaR_{1000})} - 1 \right]$, where $VaR_{(k)}$ denotes the k -th order statistic. From **Figure 1** it is clear that the spread, as measured by the 90% band, declines with increasing lambda, but increases with increasing EVI.

In principle, infinitely many repetitions are required to get the exact true VaR. The large number of simulation repetitions involved in the MC approaches above motivates the use of other numerical methods such as Panjer recursion, methods based on fast Fourier transforms [5] and the single loss approximation (SLA) method (see e.g. [6]). For a detailed comparison of numerical approximation

η	α	τ	EVI
1.00	5.00	0.60	0.33
1.00	2.00	0.60	0.83
1.00	1.00	1.00	1.00
1.00	1.50	0.50	1.33
1.00	0.30	1.80	1.85
1.00	0.17	2.50	2.35

Table 1.
Parameter sets of Burr distribution.

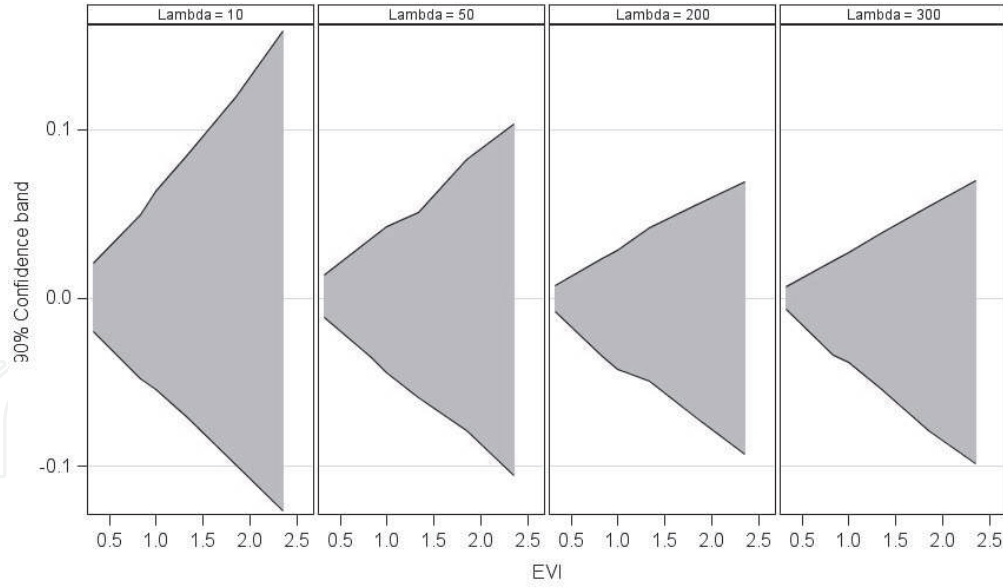


Figure 1.
Variation obtained in the VaR estimates for different values of EVI and frequency.

methods, the interested reader is referred to [7]. The SLA has become very popular in the financial industry due to its simplicity and can be stated as follows: If T is the true underlying severity distribution function of the individual losses and λ the true annual frequency then the $100(1 - \gamma)\%$ VaR of the compound loss distribution may be approximated by $T^{-1}(1 - \gamma/\lambda)$ or, as modified by [8] for large λ , by $T^{-1}(1 - \gamma/\lambda) + \lambda\mu$, where μ is the finite mean of the true underlying severity distribution. The first order approximation by [6]

$$CoP^{-1}(1 - \gamma) \approx T^{-1}(1 - \gamma/\lambda), \quad (1)$$

states that the $100(1 - \gamma)\%$ VaR of the aggregate loss distribution may be approximated by the $100(1 - \gamma/\lambda)\%$ VaR of the severity distribution, if the latter is part of the sub-exponential class of distributions. This follows from a theorem from extreme value theory (EVT) which states that $P\left(A = \sum_{n=1}^N X_n > x\right) \approx P(\max\{X_1, \dots, X_N\} > x)$ as $x \rightarrow \infty$ (see e.g. [9]). The result is quite remarkable in that a quantile of the aggregate loss distribution may be approximated by a more extreme quantile (if $\lambda > 1$) of the underlying severity distribution. EVT is all about modelling extremal events and is especially concerned about modelling the tail of a distribution (see e.g. [10]), i.e. that part of the distribution we are most interested in. Bearing this in mind we might consider modelling the body and tail of the severity distribution separately as follows.

Let q be a quantile of the severity distribution T . We use q as a threshold that splice T in such a way that the interval below q is the expected part and the interval above q the unexpected part of the severity distribution. Define two distribution functions

$$\begin{aligned} T_e(x) &= T(x)/T(q) \text{ for } x \leq q \text{ and} \\ T_u(x) &= [T(x) - T(q)]/[1 - T(q)] \text{ for } x > q, \end{aligned} \quad (2)$$

i.e. $T_e(x)$ is the conditional distribution function of a random loss $X \sim T$ given that $X \leq q$ and $T_u(x)$ is the conditional distribution function given that $X > q$.

Note that we then have the identity

$$T(x) = T(q)T_e(x) + [1 - T(q)]T_u(x) \text{ for all } x. \quad (3)$$

This identity represents $T(x)$ as a mixture of the two conditional distributions. Instead of modelling $T(x)$ with a class of distributions $F(x, \theta)$ we may now consider modelling $T_e(x)$ with $F_e(x, \theta)$ and $T_u(x)$, with $F_u(x, \theta)$. Borrowing from EVT a popular choice for $F_u(x, \theta)$ could be the generalised Pareto distribution (GPD), whilst a host of choices are available for $F_e(x, \theta)$, the obvious being the empirical distribution. Note that the Pickands-Balkema-de Haan limit theorem (see e.g. [11]), states that the conditional tail of all distributions in the domain of attraction of the Generalised Extreme Value distribution (GEV), tends to a GPD distribution. The distributions in the domain of attraction of the GEV are a wide class of distributions, which includes most distributions of interest to us. Although one could consider alternative distributions to the GPD for modelling the tail of a severity distribution, this theorem, and the limiting conditions that we are interested in, suggest that the GPD is a good choice. In the fourth section (Estimating VaR) we will discuss this in more detail.

3. Historical data and scenario modelling

It is practice in operational risk management to use different data sources for modelling future losses. Banks have been collecting their own data, but realistically, most banks only have between five and ten years of reliable loss data. To address this shortcoming, loss data from external sources and scenario data can be used by banks in addition to their own internal loss data and controls [12]. Certain external loss databases exist, including publicly available data, insurance data and consortium data. The process of incorporating data from external sources requires due consideration because of biases in the external data. One method of combining operational losses collected from various banks of different sizes and loss reporting thresholds, is discussed in [13]. In the remainder of our discussion we will only refer to historical data, which may be a combination of internal and external loss data.

Three types of scenario assessments are also suggested to improve the estimation of the severity distribution, namely the individual scenario approach, the interval approach, and the percentile approach. In the remainder of the chapter we discuss the percentile approach as we believe it is the most practical of the existing approaches available in the literature [4]. That being said, it should be noted that probability assessments by experts are notoriously difficult and unreliable as discussed in [14]. We mentioned previously that it is often an extreme quantile of the aggregate loss distribution that is of interest. In the case of operational risk, the regulator requires that the one-in-a-thousand-year quantile of this distribution be estimated, in other words the aggregate loss level that will be exceeded once in a thousand years. Considering that banks' only have limited historical data available, i.e. maximum of ten years of internal data, the estimation of such a quantile, using historical data only, is a near impossible task. So modellers have suggested the use of scenarios and experts' assessments thereof.

We advocate the use of the so-called 1-in- c year scenario approach as discussed in [4]. In the 1-in- c years scenario approach, the experts are asked to answer the question: 'What loss level q_c is expected to be exceeded once every c years?'. Popular choices for c vary between 5 and 100 and often 3 values for c are used. As an example, the bank alluded to at the start of this chapter, used $c = 7, 20$ and 100 and motivated the first choice as the number of years of reliable historical data available to them. In this case the largest loss in the historical data may serve as a guide for

choosing q_7 since this loss level has been reached once in 7 years. If the experts judge that the future will be better than the past, they may want to provide a lower assessment for q_7 than the largest loss experienced so far. If they foresee deterioration, they may judge that a higher assessment is more appropriate. The other choices of c are selected in order to obtain a scenario spread within the range that one can expect reasonable improvement in accuracy from the experts' inputs. Of course, the choice of $c = 100$ may be questionable because judgements on a 1-in-100 years loss level are likely to fall outside many of the experts' experience. In the banking environment, they may also take additional guidance from external data of similar banks which in effect amplifies the number of years for which historical data are available. It is argued that this is an essential input into scenario analysis [12]. Of course requiring that the other banks are similar to the bank in question may be a difficult issue and the scaling of external data in an effort to make it comparable to the bank's own internal data raises further problems (see e.g. [15]). We will not dwell on this issue here and henceforth assume that we do have the 1-in- c years scenario assessments for a range of c -values, but have to keep in mind that subjective elements may have affected the reliability of the assessments.

If the annual loss frequency is $Poi(\lambda)$ distributed and the true underlying severity distribution is T , and if the experts are of oracle quality in the sense of actually knowing λ and T , then the assessments provided should be

$$q_c = T^{-1}\left(1 - \frac{1}{c\lambda}\right). \quad (4)$$

To see this, let N_c denote the number of loss events experienced in c years and let M_c denote the number of these that are actually greater than q_c . Then $N_c \sim Poi(c\lambda)$ and the conditional distribution of M_c given N_c is binomial with parameters N_c and $1 - p_c = P(X \geq q_c) = 1 - T(q_c)$ with $X \sim T$ and $p_c = T(q_c) = 1 - \frac{1}{c\lambda}$. Therefore $EM_c = E[E(M_c|N_c)] = E[N_c(1 - p_c)] = c\lambda(1 - T(q_c))$. Requiring that $EM_c = 1$, yields (4).

As illustration of the complexity of the experts' task, take $\lambda = 50$ then $q_7 = T^{-1}(0.99714)$, $q_{20} = T^{-1}(0.999)$ and $q_{100} = T^{-1}(0.9998)$ which implies that the quantiles that have to be estimated are very extreme.

Returning to the SLA i.e. $CoP^{-1}(1 - \gamma) \approx T^{-1}(1 - \gamma/\lambda)$, and by taking $\gamma = 0.001$, which implies $c = 1000$, we could ask the oracle the question 'What loss level q_{1000} is expected to be exceeded once every 1000 years?'. The oracle will then produce an answer that can be used directly as an approximation for the 99.9% VaR of the aggregate loss distribution. Of course, the experts we are dealing with are not of oracle quality.

In the light of the above arguments one has to take in consideration: (a) the SLA gives only an approximation to the VaR we are trying to estimate, and (b) experts are very unlikely to have the experience or the information at their disposal to assess a 1-in-1000 year event reliably. One can realistically only expect them to assess events occurring more frequently such as once in 30 years.

Returning to the oracle's answer in (4), the expert has to consider both the true severity distribution and the annual frequency when an assessment is provided. In order to simplify the task of the expert, consider the mixed model in (3) discussed in the previous section. This model will assist us in formulating an easier question for the expert to answer. Note that the oracle's answer to the question in the previous setting can be stated as $T(q_c) = 1 - \frac{1}{c\lambda}$ (from (4)) and therefore depends on the annual frequency. However using the definition of T_u and taking $q = q_b$, $b < c$; it follows that $T_u(q_c) = 1 - \frac{b}{c}$ which does not depend on the annual frequency. This

fact that $q_c = T^{-1}(1 - \frac{1}{c\lambda}) = T_u^{-1}(1 - \frac{b}{c})$ has interesting suggestions about the formulation of the basic question of the 1-in- c years approach. For example, if we take $b = 1$ then q_1 would be the experts' answer to the question 'What loss level is expected to be exceeded once annually?'. Unless we are dealing with only rare loss events, a reasonably accurate assessment of q_1 should be possible. Then $T_u(q_c) = 1 - 1/c$ or $1 - T_u(q_c) = 1/c$. Keeping in mind the conditional probability meaning of T_u this tells us that q_c would be the answer to the question: 'Amongst those losses that are larger than q_1 , what level is expected to be exceeded only once in c years?'. Conditioning on the losses larger than q_1 has the effect that the annual frequency of all losses drops out of consideration when an answer is sought. In the remainder of the chapter we will assume that this question is posed to the experts to make their assessments.

4. Estimating VaR

Suppose we have available a years of historical loss data x_1, x_2, \dots, x_K and scenario assessments $\tilde{q}_7, \tilde{q}_{20}$ and \tilde{q}_{100} provided by the experts. In the previous sections two modelling options have been suggested for modelling the true severity distribution T and a third will follow below. The estimation of the 99.9% VaR of the aggregate loss distribution is of interest and we will consider three approaches to estimate it, namely the naïve approach, the GPD approach and Venter's approach. The naïve approach will make use of historical data only, the GPD approach (which is based on the mixed model formulation) and Venter's approach will make use of both historical data and scenario assessments. Below we demonstrate that, as far as estimating VaR is concerned, that Venter's approach is preferred to the GPD and naïve approaches.

4.1 Naïve approach

Assume that we have available only historical data and that we collected the loss severities of a total of K loss events spread over a years and denote these observed or historical losses by x_1, \dots, x_K . Then the annual frequency is estimated by $\hat{\lambda} = K/a$. Let $F(x; \theta)$ denote a suitable family of distributions to model the true loss severity distribution T . The fitted distribution is denoted by $F(x; \hat{\theta})$, with $\hat{\theta}$ denoting the (maximum likelihood) estimate of the parameter(s) θ . In order to estimate VaR a small adjustment of the Monte Carlo approximation approach, discussed earlier, is necessary.

4.1.1 Naïve VaR estimation algorithm

- i. Generate N from the Poisson distribution with parameter $\hat{\lambda}$;
- ii. Generate $X_1, \dots, X_N \sim iid F(x; \hat{\theta})$ calculate $A = \sum_{n=1}^N X_n$;
- iii. Repeat i and ii I times independently to obtain $A_i, i = 1, 2, \dots, I$. Then the 99.9% VaR is estimated by $A_{([0.999 * I] + 1)}$ where $A_{(i)}$ denotes the i -th order statistic and $[k]$ the largest integer contained in k .

4.1.2 Remarks

The estimation of VaR using the above-mentioned naïve approach has been discussed in several books and papers (see e.g. [11]). [16] stated that heavy-tailed

data sets are hard to model and require much caution when interpreting the resulting VaR estimates. For example, a single extreme loss can cause drastic changes in the estimate of the means and variances of severity distributions even if a large amount of loss data is available. Annual aggregate losses will typically be driven by the value of the most extreme losses and the high quantiles of the aggregate annual loss distribution are primarily determined by the high quantiles of the severity distributions containing the extreme losses. Two different severity distributions for modelling the individual losses may both fit the data well in terms of goodness-of-fit statistics yet may provide capital estimates which may differ by billions. Certain deficiencies of the naïve estimation approach, in particular, the estimation of the severity distribution and the subsequent estimation of an extreme VaR of the aggregate loss distribution, are highlighted in [15].

In **Figure 2** below we used the naïve approach to illustrate the effect of some of the above-mentioned claims. In **Figure 2(a)** we assumed a Burr distribution, i.e. $T_Burr(1, 0.6, 2)$, as our true underlying severity distribution. In the top panel we show the distribution function and in the middle the log of 1 minus the distribution

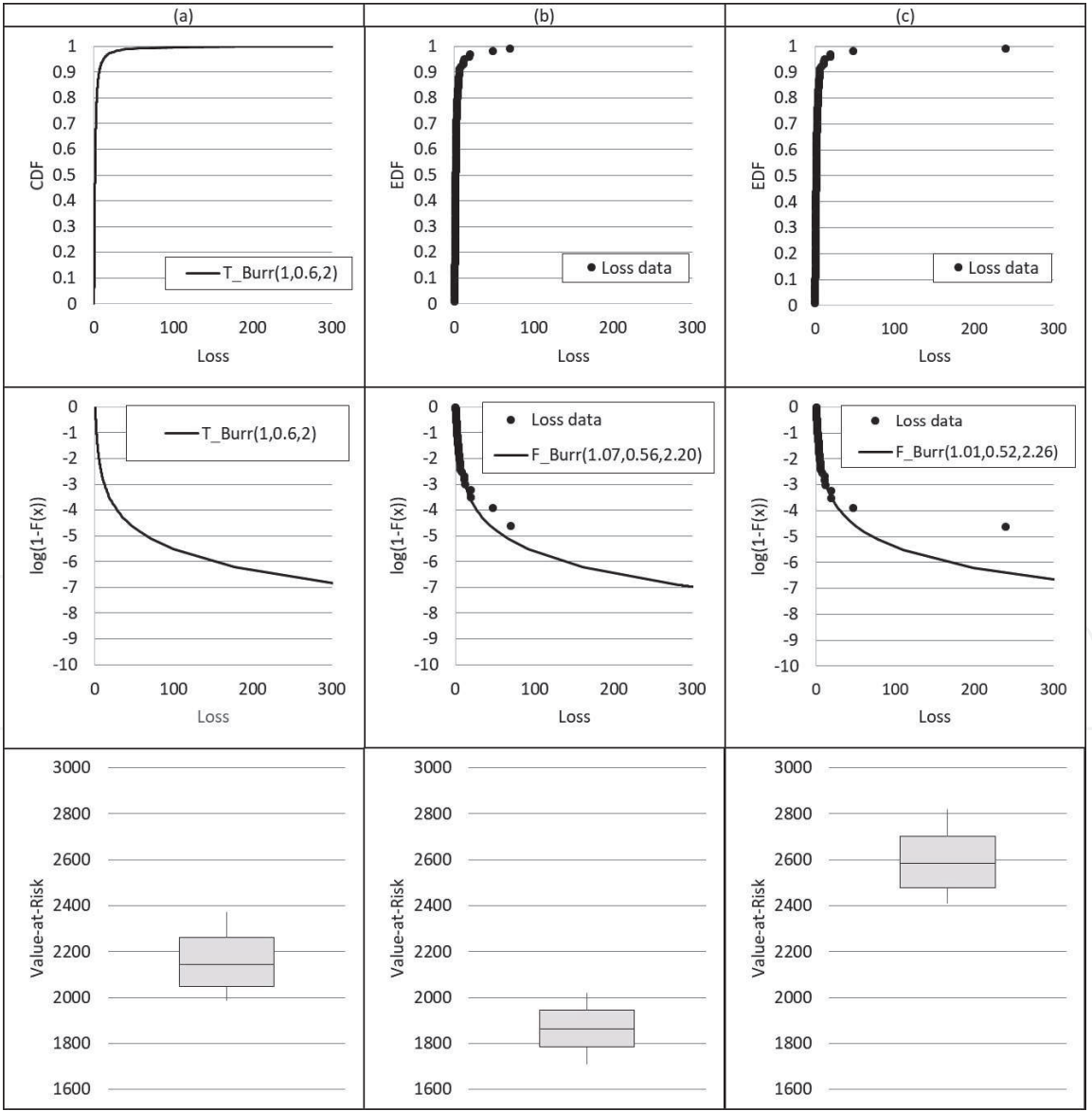


Figure 2. Illustration of the effects of VaR estimation using the naïve approach. (a) True Burr distribution, $T_Burr(1, 0.6, 2)$, (b) simulated observations from the $T_Burr(1, 0.6, 2)$ distribution with fitted distribution $F_Burr(1.07, 0.56, 2.2)$, (c) augmented simulated observations with fitted distribution $F_Burr(1.01, 0.52, 2.26)$.

function. This gives us more accentuated view of the tail of the distribution. Then in the bottom panel the Monte Carlo results of the VaR approximations are given by means of a box plot using the 5% and 95% percentiles for the box. As before, one million simulations were used to approximate VaR and the VaR calculations were repeated a 1000 times. In **Figure 2(b)** we assume $\lambda = 10$, $a = 10$ and generated 100 observations from the T_Burr(1, 0.6, 2) distribution. The observations generated is plotted in the top panel and in the middle panel the fitted distribution and the maximum likelihood estimates of the parameters are depicted as F_Burr(1.07, 0.56, 2.2). In the bottom panel the results of the VaR estimates using the naïve approach is provided. Note how the distribution of the VaR estimates differ from those obtained using the true underlying severity distribution. Of course, sampling error is present, and the generation of another sample will result in a different box plot. Let us illustrate this by studying the effect of extreme observations. In order to do this, we moved the maximum value further into the tail of the distribution and repeat the fitting process. The data set is depicted in the top panel of **Figure 2(c)** and the fitted distribution in the middle as F_Burr(1.01, 0.52, 2.26). Again, the resulting VaR estimates are shown in the bottom panel. In this case the introduction of the extreme loss has a profound boosting effect on the resulting VaR estimates.

In practice, and due to imprecise loss definitions, risk managers may incorrectly group two losses into one extreme loss that has a profound boosting effect on VaR estimates. In the light of this, it is important that the manager is aware of the process generating the data and the importance of clear definitions of loss events.

4.2 The GPD approach

This modelling approach is based on the mixed model formulation (3). As before, we have available a years of historical loss data x_1, x_2, \dots, x_K and scenario assessments $\tilde{q}_7, \tilde{q}_{20}$ and \tilde{q}_{100} . Then the annual frequency λ can again be estimated as $\hat{\lambda} = K/a$. Next b and the threshold $q = q_b$ must be specified. One possibility is to take b as the smallest of the scenario c -year multiples and to estimate q_b as the corresponding smallest of the scenario assessments \tilde{q}_b provided by the experts, in this case \tilde{q}_7 . $T_e(x)$ can be estimated by fitting a parametric family $F_e(x, \theta)$ (such as the Burr) to the data x_1, x_2, \dots, x_K or by calculating the empirical distribution and then conditioning it to the interval $(0, \tilde{q}_b]$. Either of these estimates is a reasonable choice especially if K is large and the parametric family is well chosen. Whichever estimate we use, denote it by $\tilde{F}_e(x)$. For the sake of future notational consistency, we shall also put tildes on all estimates of distribution functions which involve use of the scenario assessments.

Next, $F_u(x)$ can be modelled by the $GPD(x; \sigma, \xi, q_b)$ distribution. See Appendix A for the characteristics of this distribution. For ease of explanation, suppose we have actual scenario assessments $\tilde{q}_7, \tilde{q}_{20}$ and \tilde{q}_{100} and thus take $b = 7$ and estimate q_b by \tilde{q}_7 . Substituting these scenario assessments into $F_u(q_c) = 1 - \frac{b}{c}$; with $b = 7$, $c = 20, 100$ yields two equations.

$$F_u(\tilde{q}_{20}) = GPD(\tilde{q}_{20}; \sigma, \xi, \tilde{q}_7) = 0.65 \text{ and } F_u(\tilde{q}_{100}) = GPD(\tilde{q}_{100}; \sigma, \xi, \tilde{q}_7) = 0.93 \quad (5)$$

that can be solved to obtain estimates $\tilde{\sigma}$ and $\tilde{\xi}$ of the parameters σ and ξ in the GPD that are based on the scenario assessments. Some algebra shows that a solution exists only if $\frac{\tilde{q}_{100} - \tilde{q}_7}{\tilde{q}_{20} - \tilde{q}_7} > 2.533$. This fact should be borne in mind when the experts do their assessments.

With more than three scenario assessments, fitting techniques can be based on (5) which links the quantiles of the GPD to the scenario assessments. An example would be to minimise $\sum_c |GPD(\tilde{q}_c; \sigma, \xi, \tilde{q}_7) - (1 - b/c)|$. Other possibilities include a weighted version of the sum of deviations in this expression or deviation measures comparing the GPD quantiles directly to the q_c assessments. Whichever route we follow, we denote the final estimate of $F_u(x)$ by $\tilde{F}_u(x)$. All these ingredients can now be substituted into (3) to yield the estimate $\tilde{F}(x)$ of $T(x)$, namely

$$\hat{\lambda}\tilde{F}(x) = \left(\hat{\lambda} - \frac{1}{7}\right)\tilde{F}_e(x) + \frac{1}{7}\tilde{F}_u(x). \quad (6)$$

Returning now to practical use of Eq. (6), the algorithm below summarises the integration of the historical data with the 1-in- c years scenarios following the MC approach.

4.2.1 GPD VaR estimation algorithm

- i. Generate $N_e \sim Poi(\hat{\lambda} - \frac{1}{7})$ and $N_u \sim Poi(\frac{1}{7})$;
- ii. Generate $X_1, \dots, X_{N_e} \sim iid \tilde{F}_e$ and $X_{N_e+1}, \dots, X_{N_e+N_u} \sim iid \tilde{F}_u$ and calculate $A = \sum_{n=1}^N X_n$ where $N = N_u + N_e$. Using the identity above it easily follows that A is distributed as a random sum of N i.i.d. losses from \tilde{F} .
- iii. Repeat i and ii I times independently to obtain $A_i, i = 1, 2, \dots, I$ and estimate the 99.9% VaR by the corresponding empirical quantile of these A_i 's as before.

4.2.2 Remarks

When using the GPD 1-in- c years integration approach to model the severity distribution, we realised that the 99.9% VaR of the aggregate distribution is almost exclusively determined by the scenario assessments and their reliability greatly affects the reliability of the VaR estimate. The SLA supports this conclusion. As noted above, the SLA implies that we need to estimate $q_{1000} = T^{-1}(1 - \frac{1}{1000\lambda})$ and its estimate would be $\hat{q}_{1000} = GPD^{-1}\left(\frac{(1 - \frac{1}{1000\lambda})}{1 - (1 - \frac{1}{7\lambda})}, \tilde{\sigma}, \tilde{\xi}, \tilde{q}_b\right)$. Therefore 99.9% VaR largely depends on the GPD fitted with the scenario assessments. In **Figure 3** below we depict the VaR estimation results by fitting \tilde{F}_e assuming a Burr distribution and \tilde{F}_u assuming a GPD. The top panel in **Figure 3(a)** depicts the tail behaviour of the true severity distribution which is assumed as a Burr and denoted as T_Burr(1,0.6,2). Using the VaR approximation technique discussed in the second section (Approximating VaR) and assuming $\lambda = 10$, $I = 1\,000\,000$ and 1000 repetitions, the VaR approximations are depicted in the bottom panel in the form of a box plot as before. Assuming that we were supplied with quantile assessments by the oracle we use the two samples discussed in **Figure 2** and apply the GDP approach. The results are displayed in **Figure 3(b)** and **(c)** below.

The GPD fit to the oracle quantiles produce similar box plots, which in turn is very similar to the box plot of the VaR approximations. Clearly the fitted Burr has little effect on the VaR estimates. The VaR estimates obtained through the GPD approach is clearly dominated by the oracle quantiles. Of course, if the assessments

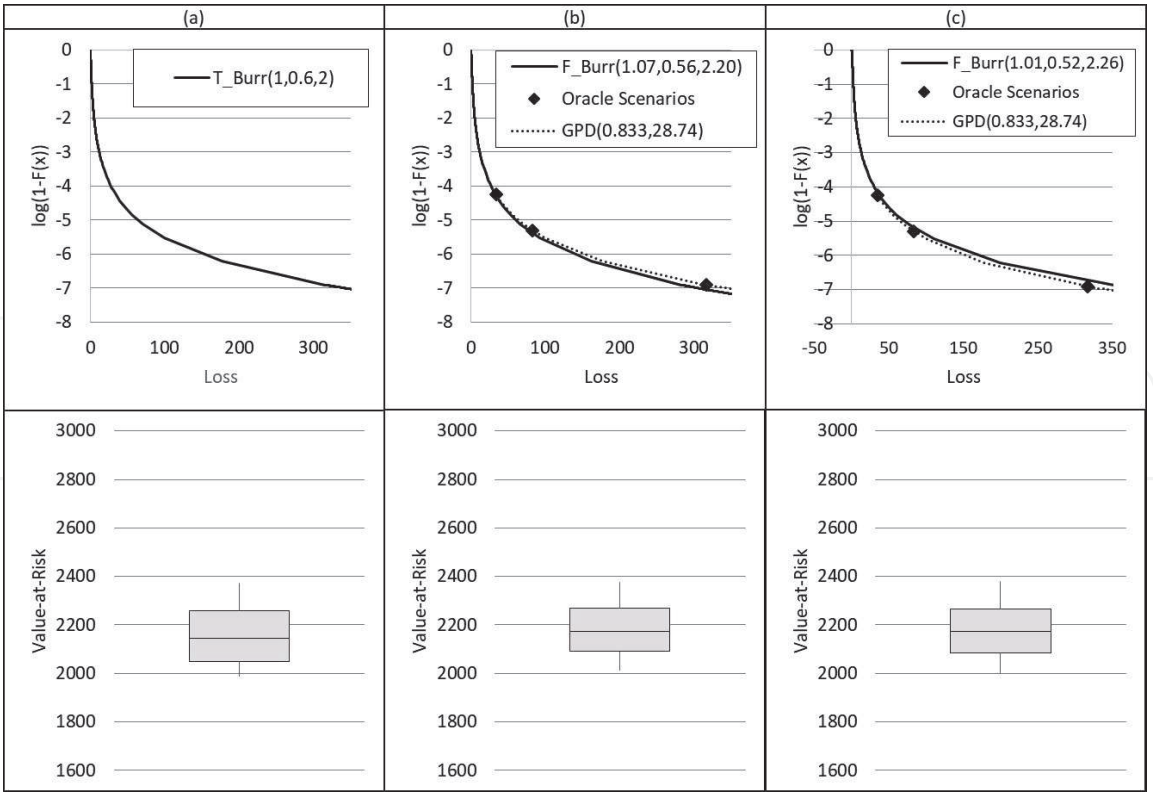


Figure 3. Illustration of VaR estimates obtained from a GPD fit on the oracle quantiles. (a) True Burr distribution, $T_Burr(1, 0.6, 2)$, (b) fitted distribution $F_Burr(1.07, 0.56, 2.2)$ on simulated data, (c) fitted distribution $F_Burr(1.01, 0.52, 2.26)$ on augmented simulated data.

are supplied by experts and not oracles the results would differ significantly. This is illustrated when we compare the GPD with Venter’s approach.

The challenge is therefore to find a way of integrating the historical data and scenario assessments such that both sets of information are adequately utilised in the process. In particular, it would be beneficial to have measures indicating whether the experts’ scenario assessments are in line with the observed historical data, and if not, to require them to produce reasons why their assessments are so different. Below we describe Venter’s estimation method that will meet these aims.

4.3 Venter’s approach

A colleague, Hennie Venter suggested that, given the quantiles q_7, q_{20}, q_{100} ; one may write the distribution function T as follows:

$$T(x) = \begin{cases} \frac{p_7}{T(q_7)} T(x) & \text{for } x \leq q_7 \\ p_7 + \frac{p_{20} - p_7}{T(q_{20}) - T(q_7)} [T(x) - T(q_7)] & \text{for } q_7 < x \leq q_{20} \\ p_{20} + \frac{p_{100} - p_{20}}{T(q_{100}) - T(q_{20})} [T(x) - T(q_{20})] & \text{for } q_{20} < x \leq q_{100} \\ p_{100} + \frac{1 - p_{100}}{1 - T(q_{100})} [T(x) - T(q_{100})] & \text{for } q_{100} < x < \infty. \end{cases} \tag{7}$$

Again $T(q_c) = p_c = 1 - \frac{1}{c\lambda}$ and it should be clear that the expressions on the right reduces to $T(x)$. Also, the definition of $T(x)$ could easily be extended for more quantiles. Given the previous discussion we can model $T(x)$ by $F(x, \theta)$ and estimate it by $F(x, \hat{\theta})$ using the historical data and maximum likelihood and estimate the

annual frequency by $\hat{\lambda} = K/a$. Given scenario assessments $\tilde{q}_7, \tilde{q}_{20}$ and \tilde{q}_{100} , then $T(q_c)$ can be estimated by $F(\tilde{q}_c, \hat{\theta})$ and p_c by $\hat{p}_c = 1 - \frac{1}{c\hat{\lambda}}$. The estimated ratios are then defined by

$$\begin{aligned} R(7) &= \frac{\hat{p}_7}{F(\tilde{q}_7; \hat{\theta})}, R(7, 20) = \frac{\hat{p}_{20} - \hat{p}_7}{F(\tilde{q}_{20}; \hat{\theta}) - F(\tilde{q}_7; \hat{\theta})}, \\ R(20, 100) &= \frac{\hat{p}_{100} - \hat{p}_{20}}{F(\tilde{q}_{100}; \hat{\theta}) - F(\tilde{q}_{20}; \hat{\theta})} \text{ and } R(100) = \frac{1 - \hat{p}_{100}}{1 - F(\tilde{q}_{100}; \hat{\theta})} \end{aligned} \quad (8)$$

Notice that if our estimates were actually exactly equal to what they are estimating, these ratios would all be equal to 1. For example, we would then have $R(7) = p_7/T(q_7) = 1$ by (4), and similarly for the others. Our new method is to estimate the true severity distribution function T by an adjusted form of $F(x, \hat{\theta})$, then Hennie's distribution \tilde{H} is defined as follows (see de Jongh et al. 2015):

$$\tilde{H}(x) = \begin{cases} R(7)F(x; \hat{\theta}) & \text{for } x \leq \tilde{q}_7 \\ \hat{p}_7 + R(7, 20)[F(x; \hat{\theta}) - F(\tilde{q}_7; \hat{\theta})] & \text{for } \tilde{q}_7 < x \leq \tilde{q}_{20} \\ \hat{p}_{20} + R(20, 100)[F(x; \hat{\theta}) - F(\tilde{q}_{20}; \hat{\theta})] & \text{for } \tilde{q}_{20} < x \leq \tilde{q}_{100} \\ \hat{p}_{100} + R(100)[F(x; \hat{\theta}) - F(\tilde{q}_{100}; \hat{\theta})] & \text{for } \tilde{q}_{100} < x < \infty. \end{cases} \quad (9)$$

Notice again that this estimate is consistent in the sense that it actually reduces to T if all estimators are exactly equal to what they are estimating.

Also note that $\tilde{H}(\tilde{q}_7) = \hat{p}_7$, $\tilde{H}(\tilde{q}_{20}) = \hat{p}_{20}$ and $\tilde{H}(\tilde{q}_{100}) = \hat{p}_{100}$, i.e. the equivalents of $T(q_c) = p_c$ hold for the scenario assessments when estimates are substituted for the true unknowns. Hence at the estimation level the scenario assessments are consistent with the probability requirements expressed. Thus, this new estimated severity distribution estimate \tilde{H} 'believes' the scenario quantile information, but follows the distribution fitted on the historical data to the left of, within and to the right of the scenario intervals. The ratios $R(7), R(7, 20), R(20, 100)$ and $R(100)$ in (9) can be viewed as measures of agreement between the historical data and the scenario assessments and could be useful for assessing their validities and qualities. The steps required to estimate VaR using this method are as follows:

4.3.1 Venter's VaR estimation algorithm

- i. Generate $N \sim \text{Poi}(\hat{\lambda})$;
- ii. Generate $X_1, \dots, X_N \sim \text{iid } \tilde{H}$ and calculate $A = \sum_{n=1}^N X_n$;
- iii. Repeat i and ii I times independently to obtain $A_i, i = 1, 2, \dots, I$ and estimate the 99.9% VaR by the corresponding empirical quantile of these A_i 's as before.

4.3.2 Remarks

The SLA again sheds some light on this method. As noted above the SLA implies that we need to estimate $q_{1000} = T^{-1}(1 - \frac{1}{1000\lambda})$ and its estimate would be

$\hat{q}_{1000} = \tilde{H}^{-1}\left(1 - \frac{1}{1000\lambda}\right) = \tilde{H}^{-1}(\hat{p}_{1000})$. Some algebra shows that the equation $F(\hat{q}_{1000}; \hat{\theta}) = F(\tilde{q}_{100}; \hat{\theta}) + (\hat{p}_{1000} - \hat{p}_{100})/R(100)$ needs to be solved for \hat{q}_{1000} . Depending on the choice of the family of distributions $F(x, \theta)$, this may be easy (e.g. when we use the Burr family for which we have an explicit expression for its quantile function). This clearly shows that a combination of the historical data and scenario assessments is involved, and not exclusively the latter. In as much as the SLA provides an approximate to the actual VaR of the aggregate loss distribution, we may expect the same to hold for Venter's approach.

In order to illustrate the properties of this approach we assume that the true underlying severity distribution is the Burr(1.0, 0.6, 2) as before. We then construct a 'false' severity distribution as the fitted distribution to the distorted sample depicted in **Figure 2(c)**, i.e. the Burr(1.00,0.52,2.26). We refer to the true severity distribution as Burr_1 and the false one Burr_2. In **Figure 4(a)** the box plots of the VaR approximations of the two distributions are given (using the same input for the MC simulations). We then illustrate the performance of the GPD and Venter approach in two cases. The first case assumes that the correct (oracle) quantiles of Burr_1 are supplied, but that the loss data are distributed according to the false distribution Burr_2. In the second case, the quantiles of the false severity distribution are supplied, but the loss data follows the true severity distribution. The box plots of the VaR estimates are given in **Figure 4(b)** for case 1 and **Figure 4(c)** for case 2.

The behaviour of the GPD approach is as expected and the box plots corresponds to the quantiles supplied. Clearly the quantiles and not the loss data dictates the results. On the other hand, the Venter approach is affected by both the loss data and quantiles supplied. In the example studied here it seems as if the method is more affected by the quantiles than by the data. This role of the data relative to the quantiles changes positively the more loss data are supplied.

4.4 GPD and Venter model comparison

In this section we conduct a simulation study to investigate the effect on the two approaches by perturbing the quantiles of the true underlying severity distributions. We assume the six parameters sets of **Table 1** as the true underlying severity distributions and then perturb the quantiles in the following way. For each simulation run, choose three perturbation factors u_7, u_{20} and u_{100} independently and uniformly distributed over the interval $[1 - \epsilon, 1 + \epsilon]$ and then take $\tilde{q}_7 = u_7 q_7, \tilde{q}_{20} = u_{20} q_{20}$ and $\tilde{q}_{100} = u_{100} q_{100}$ but truncate these so that the final values are increasing,

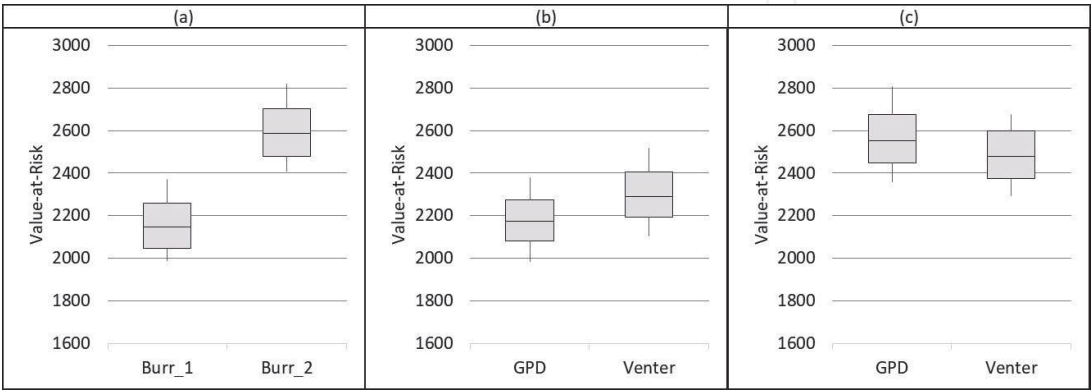


Figure 4. Comparison of VaR results for the GPD and Venter approaches. (a) Naïve approach with correct ($T_{Burr}(1, 0.6, 2)$), and false data ($F_{Burr}(1.01, 0.52, 2.26)$), (b) Case 1 with correct quantiles and false data, (c) Case 2 with false quantiles and correct data.

i.e. $\tilde{q}_7 \leq \tilde{q}_{20} \leq \tilde{q}_{100}$. Here the fraction ϵ expresses the size or extent of the possible deviations (or mistakes) inherent in the scenario assessments. If $\epsilon = 0$ then the assessments are completely correct (within the simulation context) and the experts are in effect oracles. In practice, choosing $\epsilon > 0$ is more realistic, but how large the choice should be is not clear and we therefore vary ϵ over a range of values. We chose the values 0, 0.1, 0.2, 0.3 and 0.4 for this purpose in the results below. Choosing the perturbation factors to be uniformly distributed over the interval $[1 - \epsilon, 1 + \epsilon]$ implies that on average they have the value 1, i.e. the scenario assessments are about unbiased. This may not be realistic and other choices are possible, e.g. we could mimic a pessimistic scenario maker by taking the perturbations to be distributed on the interval $[1, 1 + \epsilon]$ and an optimistic scenario maker by taking them on the interval $[1 - \epsilon, 1]$.

For each combination of parameters of the assumed true underlying Poisson frequency and Burr severity distributions and for each choice of the perturbation size parameter ϵ the following steps are followed:

- i. Use the VaR approximation algorithm in the second section to determine the 99.9% VaR for the Burr Type XII with the current choice of parameters. Note that the value obtained here approximately equals the true 99.9% VaR. We refer to this value as the approximately true (AT) VaR.
- ii. Generate a data set of historical losses, i.e. generate $K \sim Poi(7\lambda)$ and then generate $x_1, x_2, \dots, x_K \sim iid$ Burr Type XII with the current choice of parameters. Here the family $F(x, \theta)$ is chosen as the Burr Type XII but it is refitted to the generated historical data to estimate the parameters as required.
- iii. Add to the historical losses three scenarios $\tilde{q}_7, \tilde{q}_{20}, \tilde{q}_{100}$ generated by the quantile perturbation scheme explained above. Estimate the 99.9% VaR using the GPD approach.
- iv. Using the historical losses and the three scenarios of item iii), calculate the severity distribution estimate \tilde{H} and apply Venter's approach to estimate the 99.9% VaR.
- v. Repeat items i–iv 1000 times and then summarise and compare the resulting VaR estimates.

Because we are generally dealing with positively skewed data here, we shall use the median as the principal summary measure. Denote the median of the 1000 AT values by $MedAT$. Then we construct 90% VaR bands as before for the 1000 repeated GPD and Venter VaR estimates, i.e. $\left[\frac{VaR_{(51)}}{MedAT} - 1, \frac{VaR_{(951)}}{MedAT} - 1 \right]$. The results are given in **Figure 5**. Note that light grey represents the GPD band and dark grey the Venter band, whilst the overlap between the two bands are even darker.

From **Figure 5**, we make the following observations:

For small frequencies ($\lambda \leq 10$) the GPD approach outperforms the Venter approach, except for short tailed severity distributions and higher quantile perturbations. When the annual frequency is high ($\lambda \geq 50$) and for moderate to high quantile perturbations ($\epsilon \geq 0.2$) the Venter approach is superior, and more so for higher λ and ϵ . Even for small quantile perturbations ($\epsilon = 0.1$) and high annual frequencies ($\lambda \geq 50$) the Venter approach performs reasonable when compared to the GPD.

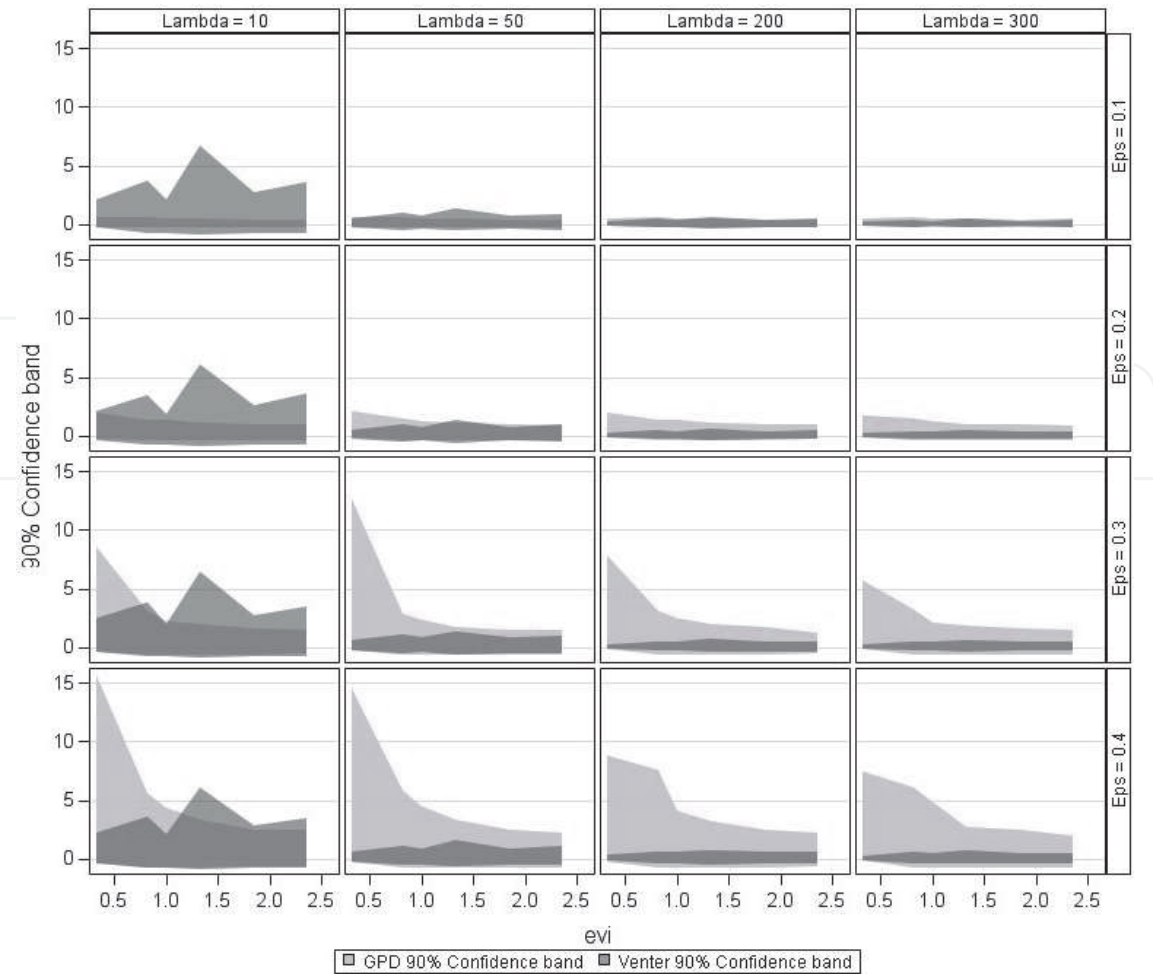


Figure 5.
VaR bands for different Burr parameter sets and frequency combinations.

The above information suggest that provided enough loss data is available the Venter approach is the best choice to work.

5. Implementation recommendations

As stated in the introduction to this chapter, Venter’s method has been implemented by major international banks and approved by the local regulator. Based on this experience, we can share the following implementation guidelines:

- i. Study the loss data carefully with respect to the procedures used to collect the data. Focus should be on the largest losses and one has to establish whether these losses were recorded and classified correctly according to the definitions used.
- ii. Experts should be presented with an estimate of q_1 (based on the loss data) and then should answer the question ‘Amongst those losses that are larger than q_1 what level is expected to be exceeded only once in c years?’ where $c = 7, 20, 100$.
- iii. The assessments by the expert should be checked with the condition $\frac{\tilde{q}_{100}-\tilde{q}_7}{\tilde{q}_{20}-\tilde{q}_7} > 2.533$. This bring realism as far as the ratios between the assessments are concerned.

- iv. The loss data may be fitted by a wide class of severity distributions. We used SAS PROC SEVERITY in order to identify the five best fitting distributions.
- v. Calculate the ratios $R(7)$, $R(7, 20)$, $R(20, 100)$ and $R(100)$ of the best fitting distributions obtained above and then select the best distribution based on the ratios. Although this is a subjective selection it will lead to more realistic choices.
- vi. For the best fitting distribution, present the ratios that deviate significantly from one to the experts for possible re-assessment. If new assessments are provided, repeat guidelines iii to v once or twice.
- vii. Different data sources should be considered. The approaches discussed above assumes one unified dataset for the historical data source. In practice different datasets are included for example internal, external and mixed where the latter is scaled. Estimates of q_1 and q_7 based on these different datasets should inform the scenario process.
- viii. Guideline vi may also be repeated on appropriate mixed (scaled) data sets to select the best distribution type.

6. Some further practical considerations

Data Scaling. It is practice in operational risk management to use different data sources for modelling future losses. Banks have been collecting their own data, but realistically, most banks only have between five and ten years of reliable loss data. To address this shortcoming, loss data from external sources can be used by banks in addition to their own internal loss data and controls. External loss data comprises operational risk losses experienced by third parties, including publicly available data, insurance data and consortium data. [16] investigate whether the size of operational risk losses is correlated with geographical region and firm size. They use a quantile matching algorithm to address statistical issues that arise when estimating loss scaling models when subjecting the data to a loss reporting threshold. [13] uses regression analysis based on the GAMLSS (generalised additive models for location scale and shape) framework to model the scaling properties. The severity of operational losses using the extreme value theory is used to account for the reporting bias of the external data losses.

No historical data available. In the event of having insufficient historical data available, the GPD approach as discussed above may be used. $T_e(x)$ in (2) can be estimated by a right truncated distribution, e.g. scaled beta, Pareto type II, etc. fitted to an expected loss scenario and q_7 . In this case the expert should also provide a scenario for the expected loss $EL = E[T|X \leq q_7]$. $T_u(x)$ can be estimated by a GPD distribution as discussed in the GPD approach.

Aggregation. To capture dependencies of potential operational risk losses across business lines or event types, the notion of copulas may be used (see [15]). Such dependencies may result from business cycles, bank-specific factors, or cross-dependence of large events. Banks employing more granular modelling approaches may incorporate a dependence structure, using copulas to aggregate operational risk losses across business lines and/or event types for which separate operational risk models are used.

7. Conclusion

In this chapter, we motivated the use of Venter's approach whereby the severity distribution may be estimated using historical data and experts' scenario assessments jointly. The way in which historical data and scenario assessments are integrated incorporates measures of agreement between these data sources, which can be used to evaluate the quality of both. This method has been implemented by major international banks and we included guidelines for its practical implementation. As far as future research is concerned, we are investigating the effectiveness of using the ratios in assisting the experts with their assessments. Also, we are testing the effect of replacing q_{100} with q_{50} in the assessment process.

A. Appendix A

A.1 The generalised Pareto distribution (GPD)

The GPD given by

$$GPD(x; \sigma, \xi, q_b) = \begin{cases} 1 - [1 + \frac{\xi}{\sigma}(x - q_b)]^{-\frac{1}{\xi}} & \xi > 0 \\ 1 - \exp\left(-\frac{x - q_b}{\sigma}\right) & \xi = 0, \end{cases} \quad (10)$$

with $x \geq q_b$, thus taking q_b as the so-called EVT threshold and with σ and ξ respectively scale and shape parameters. Note the Extreme Value Index (EVI) of the GPD distribution is given by $EVI = \xi$ and that heavy-tailed distributions have a positive EVI and larger EVI implies heavier tails. This follows (also) from the fact that for positive EVI the GPD distribution belongs to the Pareto-type class of distributions, having a distribution function of the form $1 - F(x) = x^{-1/\xi} \ell_F(x)$, with $\ell_F(x)$ a slowly varying function at infinity (see e.g. Embrechts et al., 1997). For Pareto-type, when the $EVI > 1$, the expected value does not exist, and when $EVI > 0.5$, the variance is infinite. Note also that the GPD distribution is regularly varying with index $-1/\xi$ and therefore belongs to the class of sub-exponential

distributions. Note that the γ -th quantile of the GPD is $q(\gamma) = GPD^{-1}(\gamma, \sigma, \xi, q_b) = \left(q_b + \frac{\sigma((1-\gamma)^{-\xi} - 1)}{\xi}\right)$ when $\xi \neq 0$ and $GPD^{-1}(\gamma, \sigma, \xi, q_b) = q_b - \sigma \ln(1 - \gamma)$ when $\xi = 0$.

A.2 The Burr distribution

The three parameter Burr type XII distribution function

$$B(x; \eta, \tau, \alpha) = 1 - (1 + (x/\eta)^\tau)^{-\alpha}, \text{ for } x > 0 \quad (11)$$

with parameters $\eta, \tau, \alpha > 0$ (see e.g. [10]). Here η is a scale parameter and τ and α shape parameters. Note the EVI of the Burr distribution is given by $EVI = \zeta = 1/\tau\alpha$ and that heavy-tailed distributions have a positive EVI and larger EVI implies heavier tails. This follows (also) from the fact that for positive EVI the Burr distribution belongs to the Pareto-type class of distributions, having a distribution function of the form $1 - F(x) = x^{-1/\zeta} \ell_F(x)$, with $\ell_F(x)$ a slowly varying function at infinity (see e.g. [9]). For Pareto-type, when the $EVI > 1$, the expected value does

not exist, and when $EVI > 0.5$, the variance is infinite. Note also that the Burr distribution is regularly varying with index $-\tau\alpha$ and therefore belongs to the class of sub-exponential distributions. Note that the γ -th quantile of the Burr distribution is $q(\gamma) = B^{-1}(\gamma; \eta, \tau, \alpha) = \eta \left((1 - \gamma)^{-1/\alpha} - 1 \right)^{1/\tau}$.

Other declarations


The authors acknowledge grants received from the National Research Foundation, the Department of Science and Technology and the Department of Trade and Industry. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors, and therefore the National Research Foundation does not accept any liability in regard to them.

Author details

Riaan de Jongh*, Helgard Raubenheimer and Mentje Gericke
Centre for Business Mathematics and Informatics, North-West University,
Potchefstroom, South Africa

*Address all correspondence to: riaan.dejongh@nwu.ac.za

IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Basel Committee on Banking Supervision. OPE Calculation of RWA for Operational Risk—OPE30 Advance Measurement Approach. 2019. Available from: https://www.bis.org/basel_framework/chapter/OPE/30.htm
- [2] Basel Committee on Banking Supervision. Basel III: Finalising Post-Crisis Reforms. 2017. Available from: <https://www.bis.org/bcbs/publ/d424.htm>
- [3] Prudential Regulation Authority, The PRA's Methodologies for Setting Pillar 2 Capital. Bank of England. 2020. Available from: <https://www.bankofengland.co.uk/-/media/boe/files/prudential-regulation/statement-of-policy/2020/the-pras-methodologies-for-setting-pillar-2a-capital-update-february-2020.pdf>
- [4] De Jongh P, De Wet T, Raubenheimer H, Venter J. Combining scenario and historical data in the loss distribution approach: A new procedure that incorporates measures of agreement between scenarios and historical data. *The Journal of Operational Risk*. 2015;**10**(1):45-76. DOI: 10.21314/JOP.2015.160
- [5] Panjer H. *Operational Risk: Modeling Analytics*. Chichester: Wiley; 2006. p. 448
- [6] Böcker K, Klüppelberg C. Operational VaR: A closed-form approximation. *Risk Magazine*. 2005;**18**(12):90-93
- [7] De Jongh P, De Wet T, Panman K, Raubenheimer H. A simulation comparison of quantile approximation techniques for compound distributions popular in operational risk. *The Journal of Operational Risk*. 2016;**11**(1):23-48. DOI: 10.21314/JOP.2016.171
- [8] Degen M. The calculation of minimum regulatory capital using single-loss approximations. *The Journal of Operational Risk*. 2010;**5**(4):3-17. DOI: 10.21314/JOP.2010.084
- [9] Embrechts P, Klüppelberg C, Mikosch T. *Modelling Extremal Events for Insurance and Finance*. Berlin, Heidelberg: Springer; 1997
- [10] Beirlant J, Goegebeur Y, Segers J, Teugels J. *Statistics of Extremes: Theory and Applications*. New Jersey: John Wiley and Sons; 2004
- [11] McNeil A, Frey R, Embrechts P. *Quantitative Risk Management: Concepts, Techniques and Tools*. Revised Edition. Princeton and Oxford: Princeton University Press; 2015
- [12] Basel Committee on Banking Supervision. *Operational Risk: Supervisory Guidelines for the Advanced Measurement Approaches*. Report 196. 2011. Available from: <https://www.bis.org/publ/bcbs196.htm>
- [13] Ganegoda A, Evans J. A scaling model for severity of operational losses using generalized additive models for location scale and shape (GAMLSS). *Annals of Actuarial Science*. 2013; **7**(1):61-100. DOI: 10.1017/S1748499512000267
- [14] Kahneman D, Slovic P, Tversky A. *Judgement under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press; 1982
- [15] Embrechts P, Hofert M. Practices and issues in operational risk modelling under Basel II. *Lithuanian Mathematical Journal*. 2011;**51**(2):180-193
- [16] Cope E, Labbi A. Operational loss scaling by exposure indicators: Evidence from the ORX database. *The Journal of Operational Risk*. 2008;**3**(4):25-45. DOI: 10.21314/JOP.2008.051